Thoughts on Critical Infrastructure Collaboration

Andrew J. Scholand
ajschol@sandia.gov

John M. Linebarger
jmlineb@sandia.gov

Mark A. Ehlen
maehlen@sandia.gov

Sandia National Laboratories[*]
Albuquerque, NM 87185-1138

ABSTRACT

In this paper, we describe what we believe to be the characteristics of the collaborations required in the domain of critical infrastructure modeling, based on our experiences to date. We adopt a knowledge management philosophy, which imposes two classes of requirements, **contextual** (*who*, *when*, and *why*), and **semantic** (*what* interactions are conducted around). We observe that infrastructure models can often engender more insight when used as the basis for a meaningful discussion between the disparate stakeholder groups (private industry, trade organizations, industry lobbying groups, etc.) than when exercised computationally.

Categories and Subject Descriptors

K.4.3 [**Computers & Society**]: Organizational Impacts—*Computer-supported collaborative work*

General Terms

Design

Keywords

Collaboration, group discussion, modeling and simulation

1. INTRODUCTION

Critical Infrastructures are the interdependent technological structures and systems that support the operation of society, such as telecommunications, transportation, energy transmission and transformation, water and wastewater, and also related social, political, emergency planning and response and economic systems. The National Infrastructure Simulation and Analysis Center (NISAC) provides analysis services addressing the impact of potential events (natural or man-made disasters, policy proposals, mitigation strategies) on these highly complex, interrelated systems. Typical NISAC questions are aspecific, open-ended, and vague, and the time frames for providing a response range from hours to months. Executive orders have defined 14 critical infrastructure categories which range from postal and shipping infrastructure to various classes of energy infrastructure (i.e. electrical power grid, petroleum distribution network, etc.); it is impossible for any one person to have sufficient expertise across this spectrum to answer the kinds of questions asked with a suitable degree of accuracy.

The information needed to address cross-infrastructure issues is also partitioned across various organizations. Approximately 85% of critical infrastructure in the U.S. is privately owned, and many of the aspects characterizing infrastructures as critical are based on economic activity in private industry. Infrastructure operators, business owners, and other members of the private sector possess the detailed data describing the current operating state of the system needed to ground domain knowledge in reality. NISAC depends heavily on collaboration with infrastructure sector experts at the national, regional, and local levels for access to this data and information about infrastructure processes and operations.

The vision for NISAC is to address the domain knowledge problem by building collaborative ventures with this diverse group of public and private entities, amalgamating widely geographically and organizationally distributed infrastructure expertise. We propose to apply the best practices of knowledge management to these interactions, so that we capture, persist, and reflect on the nature and extent of the collaborations which occur through the framework. This focus drives two classes of requirements, **contextual** requirements capturing dimensions of *who*, *when*, and *why* interactions occur, and **semantic** requirements identifying *what* interactions are conducted around. Each requirement class is addressed in a section below.

2. CONTEXTUAL REQUIREMENTS

The key advantages of computationally-mediated collaborative communication systems in supporting asynchronous communication and automatic archival still apply when viewed through a knowledge-management perspective. Asynchronous communication also lowers the cost of maintaining intersubjectivity, the explicitly shared comprehension of communicated concepts such as common meanings and metaphors. If collaboration is not tied to real time, the group's common understanding is not bound by the lowest common denominator - participants who don't yet grasp a shared concept have the latitude to research the meaning of the terms describing the concept without derailing the group dynamic. In other words, persistence, in proper form, allows later di-

gestion and retrospective commentary by individuals that can significantly enhance the value of the group's knowledge. We argue, however, that *synchronous* mechanisms, where the on-line subgroup of the overall community is interacting in real-time, are necessary for individual interpretations, understanding, and insights build upon one another to form cogent findings and conclusions under the very real time constraints of the problem domain. Synchronous communication is desirable because it enforces participation in a community, similar to the organizational behavior of 'calling a meeting' but without the colocation requirements. This is particularly significant for domain experts, who are typically highly valued for, and therefore have a number of demands on, their time. Synchronicity increases the stridence of a collaborative contribution, and for efforts where the collaborative goal is a higher priority goal than other tasks this can be beneficial at maximizing input.

We feel the KM perspective does derogate the suggested benefit of anonymous communication in collaboration. Although limited visibility of an individual's identity may be important for social/psychological factors encouraging public admission of limited knowledge or a lack of understanding, from a knowledge management perspective there is tremendous value in associating contextual information about an individual with the shared contributions that person makes. Rich meta data is needed about participants, including their educational background and disciplinary viewpoint, their experience level, their familiarity with domain, their organizational affiliation. In addition, the archival system needs to capture when questions are raised, both in terms of calendar time, and in the broader qualitative context of the stage of the analysis process unfolding during the collaboration (i.e. is the collaboration in an expanding possibilities/brainstorming phase; a narrowing/focusing down phase; an analytic or computational phase; in literature or background searching; or otherwise synthesizing or categorizing data; etc.) The objective is to create information that can later be mined to understand: *who* is asking the questions and *who* is answering them; *when* are questions raised; and *why* questions arise, in terms of what activities are being undertaken at that time.

This suggests that although identity and 'electronic persona' information may have limited visibility to the human participants in the collaboration, depending on the context of the information exchange, it needs to have global system/internal visibility for later knowledge extraction and processing. One goal of this data mining is to identify if certain classes of simulation results are triggers for questions concerning authenticity in order to create best practices for vetting simulation models before their results are presented as meaningful. Another goal is to allow later identification of experts (those with knowledge) so their views and commentary can be treated differently in various ways (extracted and compiled, searched, weighted more heavily, etc.) Group dynamics as a function of individual characteristics will also be investigated - what composition of individuals experienced the most synergistic collective problem solving? What was the composition of groups that did not prove effective at problem solving?

3. SEMANTIC REQUIREMENTS

One question not included in the list above is *what* questions were asked, and what answers were given. Collab-

oration involves sharing both data and knowledge about the data. While the latter will typically be highly unstructured, it is possible to capture a great deal of rich machine-processable information about data which are shared. Again, the end goal is to enable a knowledge extraction process to answer questions such as: What areas do experts drill into? What data do they use to verify a model presents results that match their experience? Are there commonalities in what data are looked at or looked for? What are the characteristics and distinguishing features of these data?

The programmatic means to achieve this end may in itself be worthwhile from a collaboration framework perspective - if the framework explicitly represents rich semantics in all data transfers across it, tremendous opportunities exist for software developers to automate various data-centric tasks, such as transforming exchanged data or merging new data with existing information. For example, many researchers have suggested the creation of Frequently Asked Questions (FAQs) lists as a tool to aid a collaborative community, which are of course, inherently task-based ("How do I ...?"). Semantically explicit data structures would allow the formation of Frequently Observed Objects (FOO) lists ("Who else was concerned about fixed-price options?") which may cross-cut task-based hierarchies of system functionality.

4. CONCLUSIONS

NISAC naturally depends heavily on collaboration with disparate national, regional, and local stakeholder groups for the operational data and process information necessary to create accurate and complete models, simulations, and analyses. Collaboration is also fundamental, however, to the subsequent continuous refinement of the decision support advice that can be extracted from the simulation(s). The tremendous intrinsic value of collaboration between the development environment of the laboratories and the heterogeneous stakeholder "use" environment, however, has been diminished by the transitory and ephemeral nature of the exchange of technical and conceptual knowledge in 'real-world' meetings and workshops. This paper argues for conducting this synchronous collaboration in the digital domain to enable both subsequent asynchronous evolution of shared knowledge artifacts and traditional Knowledge Management access. The extent and nature of knowledge engineering queries is probably limitless, but here we have proposed a few examples, such as "Who, exactly, is asking what questions, and can we figure out why? Do those questions indicate areas where the model needs to be improved? Are there issues the simulation needs to inform the model operator/exercisor about better?"

5. REFERENCES

[1] Hollan, J. and Stornetta, S. 1992. *Beyond being there.* In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Monterey, California, United States, May 03 - 07, 1992). P. Bauersfeld, J. Bennett, and G. Lynch, Eds. CHI '92. ACM Press, New York, NY, 119-125. DOI= http://doi.acm.org/10.1145/142750.142769